EMBO reports

scientific report

Genome-wide distribution of DNA replication origins at A+T-rich islands in *Schizosaccharomyces pombe*

Mónica Segurado, Alberto de Luis & Francisco Antequera[+]

Instituto de Microbiología Bioquímica, CSIC/Universidad de Salamanca, Edificio Departamental, Salamanca, Spain

Genome-wide analysis of replication dynamics requires the previous identification of DNA replication origins (ORIs). However, variability among the ORIs makes it difficult to predict their distribution across the genome on the basis of their sequence. We report here that ORIs in *Schizosaccharomyces pombe* coincide with discrete chromosomal A+T-rich islands of up to 1 kb long that are characterized by a distinctive A+T content that clearly differentiates them from the rest of the genome. Genome-wide analysis has enabled us to identify 384 of these regions, which predicts the position of most ORIs in the genome, as shown by functional replication analyses. A+T-rich islands occur at the mating locus, centromeres and subtelomeric regions at a density that is approximately fourfold higher than elsewhere in the genome, which suggests a link between the origin recognition complex (ORC) and transcriptional silencing in these regions. The absence of consensus elements in A+T-rich islands implies that different sequences can target the ORC to different ORIs.

EMBO reports **4**, 1048–1053 (2003)
doi:10.1038/sj.embor.7400008

INTRODUCTION

Eukaryotic DNA replication origins (ORIs) have been identified in a variety of organisms including fungi, insects and mammals, and they have been well characterized biochemically and genetically in the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. In *S. cerevisiae*, ORI regions span less than 150 bp and contain one or several copies of an 11-bp autonomous replicating sequence (ARS) consensus (Broach *et al.*, 1983; Newlon & Theis, 1993), which is essential for binding of the origin recognition complex (ORC). In addition, ORIs contain three or four partially redundant auxiliary elements the sequence and distribution of which varies between them (Theis & Newlon, 1997, 2001). Two recent approaches based on chromatin immunoprecipitation and density labelling of replication intermediates have predicted the distribution of approximately 400 putative ORIs in *S. cerevisiae* (Wyrick *et al.*, 2001; Raghuraman *et al.*, 2001).

S. pombe ORIs require a minimum length of 0.5–1 kb and do not have recognizable consensus elements. However, functional dissection analyses have identified several A+T-rich elements that frequently contain stretches of asymmetrical adenine or thymine residues. The length and number of such stretches are not conserved between different ORIs. Some of these elements are individually or collectively required for ORI activity in plasmids and in the chromosome (Clyne & Kelly, 1995; Dubey *et al.*, 1996; Okuno *et al.*, 1999; Takahashi *et al.*, 2003). A key advance in our understanding of how the ORC is targeted to ORIs in *S. pombe* was the discovery of the unique structure of its Orc4 subunit (Chuang & Kelly, 1999). The amino terminus of the SpOrc4 protein contains nine AT-hook domains that are essential for ORC binding to ORIs *in vitro* and *in vivo* (Lee *et al.*, 2001; Kong & DePamphilis, 2002). These domains recognize the structure of A+T-rich stretches through the minor groove of DNA without the requirement of a specific nucleotide sequence (Reeves & Beckerbauer, 2001). We report here the identification of discrete genomic regions up to 1 kb long that have a disctinctively high A+T composition, of which approximately 90% colocalizes with active ORIs. The properties of these A+T-rich islands may account for the specific properties of *S. pombe* ORIs and provide a frame of reference for future analyses of replication dynamics in this yeast.

RESULTS

Identification of A+T-rich islands at ORIs in *S. pombe*

All DNA replication origins identified so far in *S. pombe* are located in intergenic regions upstream from genes, although not all of these regions are competent in initiating replication (Dubey *et al.*, 1994; Okuno *et al.*, 1999; Gómez & Antequera, 1999). This observation, together with the fact that the AT-hook domains at the N terminus of the SpOrc4 protein are required for ORC binding to ORIs (Chuang & Kelly, 1999; Lee *et al.*, 2001; Kong & DePamphilis, 2002), raised the possibility that differences in the A+T content might be a determinant of ORI activity. To test this possibility, we compared base composition across 16 regions containing active genomic ORIs previously identified in our laboratory (Gómez & Antequera, 1999; Segurado *et al.*, 2002) with another 16 regions of similar length, which were also upstream

Instituto de Microbiología Bioquímica, CSIC/Universidad de Salamanca, Edificio Departamental, Campus Miguel de Unamuno, 37007 Salamanca, Spain
[+]Corresponding author. Tel: +34 923 121778; Fax: +34 923 224876;
E-mail: CpG@usal.es

Received 30 July 2003; revised 10 September 2003; accepted 12 September 2003
Published online 17 October 2003

from genes but devoid of ORI activity (table 1 of supplementary information online). Given that the shortest DNA fragments capable of conferring autonomous replication to plasmids in *S. pombe* range between 0.5 kb and 1 kb (Maundrell *et al.*, 1988; Dubey *et al.*, 1994; Clyne & Kelly, 1995), we determined base composition using sliding windows of different sizes within this interval. We found that the highest A+T content for each window was significantly higher for ORI-containing regions than for regions that replicated passively, and that the 32 regions analysed could be classified in two distinct and non-overlapping groups (Fig. 1A; see table 1 of supplementary information online for the specific value for each window). These differences were also evident when the A+T content was plotted across long genomic regions containing ORIs such as ORI 12 and ORI tug1 (Fig. 1B; Gómez & Antequera, 1999; Segurado *et al.*, 2002). On the basis of these observations, we defined A+T-rich islands as regions between 0.5 kb and 1 kb the A+T content of which was equal or above the following values for every window size (500 bp: 75%; 600 bp: 74.5%; 700 bp: 74%; 800 bp: 73%; 900 bp: 72.5%; 1,000 bp: 72%). When a particular region fulfilled the criterion for all window sizes except for one, a 0.5% reduction in the A+T content for this particular size was allowed.

To assess further the association between A+T-rich islands and ORIs, we tested whether another set of 14 ORIs identified by other authors using various approaches would also colocalize with them (green triangles in Fig. 2). Base composition analysis across regions containing *ars3001* at the rRNA gene cluster (Sánchez *et al.*, 1998), *ars3003* and *ars3002* (Dubey *et al.*, 1994), and *ars2004* (Okuno *et al.*, 1997, 1999) showed that this was indeed the case for all of them. A+T-rich islands were also found to span four ORIs in the K and L repeats of centromere II (Smith *et al.*, 1995) and eight ARS (Maundrell *et al.*, 1988), at least two of which have been shown to act as chromosomal ORIs (Dalgaard & Klar, 2001; Segurado *et al.*, 2002; table 2 of supplementary information online). Examples corresponding to *ars3003*, *ars3002*, *ars2004* and *ars1* are shown in Fig. 1C. This figure also shows that A+T-rich islands overlap with the shortest DNA fragments that are capable of maintaining full ARS activity and include the replication initiation point (RIP) in *ars2004* (Okuno *et al.*, 1997) and *ars1* (Gómez & Antequera, 1999). A+T-rich islands do not extend across the entire intergenic regions, as illustrated by island 1003 (see below), which spans only a discrete fraction of the 5-kb-long intergenic region between two divergent genes (Fig. 1C).

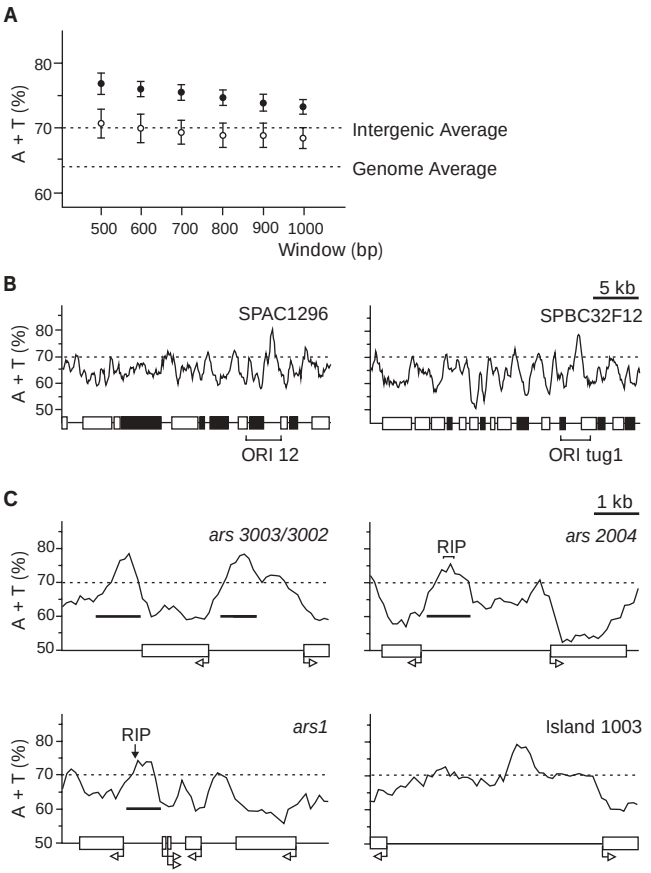

Fig. 1 | Identification of A+T-rich islands at origins of replication in *Schizosaccharomyces pombe*. (**A**) The highest A+T content of 16 origins of replication (ORI) regions and 16 non-ORI regions was determined using sliding windows of 500–1,000 bp and a step of 1 bp. Their average content, with standard deviation, is indicated. Broken lines indicate the total genomic and intergenic A+T average content (64% and 70%, respectively). (**B**) A+T content across 30 kb of cosmids SPAC1296 and SPBC32F12 with ORI 12 and ORI tug1, measured with a 500-bp window and a step of 100 bp. Black and white rectangles represent genes transcribed towards the left and the right, respectively. Brackets indicate the position of restriction fragments containing the ORIs. Broken lines indicate the intergenic average A+T content. Scale bar, 5 kb. (**C**) A+T content across 6-kb-long regions containing the autonomous replicating sequences *ars3003*, *ars3002*, *ars2004* and *ars1*, and A+T-rich island 1003, measured as in **B**. Black bars indicate the shortest DNA fragment capable of maintaining full ARS activity in each case. RIP indicates the replication initiation point in *ars1* and *ars2004*. Rectangles represent genes, and arrows indicate the direction of transcription. Scale bar, 1 kb.

Genome-wide distribution of A+T-rich islands and ORIs

The fact that the 30 previously known ORIs (tables 1 and 2 of supplementary information online) were associated with A+T-rich islands raised the possibility that the genome-wide distribution of ORIs could be predicted by localizing the position of the islands on the basis of their base composition. Therefore, we used the Artemis and EMBOSS software packages (see Methods) to search the *S. pombe* genome sequence for regions between 0.5 kb and 1 kb long the A+T content of which was higher than the boundaries described above; we identified 384 that qualified as A+T-rich islands. Their distribution along the three *S. pombe* chromosomes is represented in Fig. 2, and their genomic localization is indicated in table 3 of supplementary information online. Their average genomic frequency was one every 33 kb and, with no exceptions, they mapped at intergenic regions. A+T-rich islands were over-represented between divergent transcription units (52% observed (O), 26.6% expected (E)) and under-represented between collinear (37.7% O, 46.8% E) and convergent (10.3% O, 26.6% E) transcription units. These percentages excluded the 25 A+T-rich islands clustered in the three centromeres. A similar or even more pronounced bias was previously reported using a smaller number of cases (Gómez & Antequera, 1999). This was

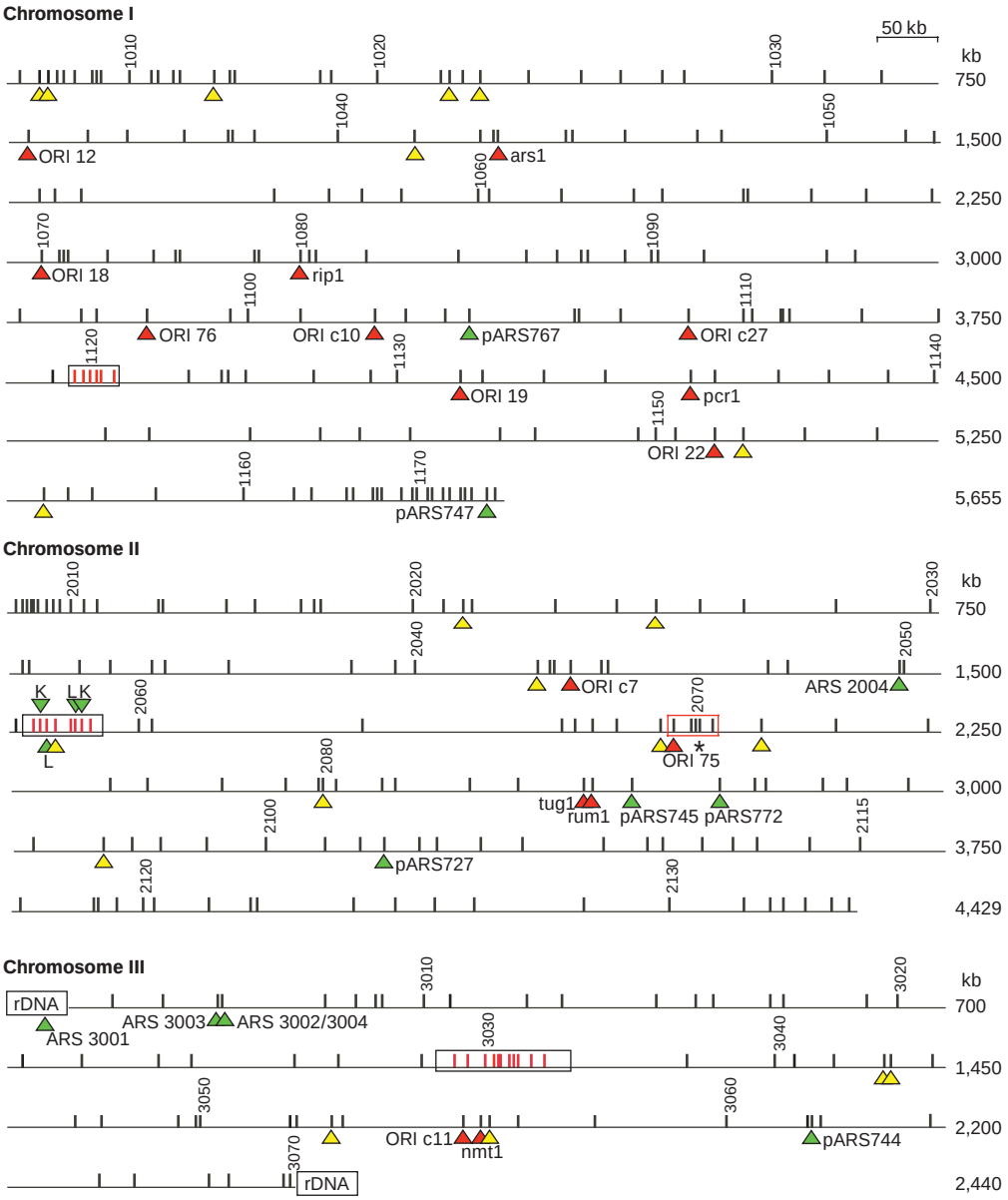

Fig. 2 | Genome-wide distribution of A+T-rich islands. Vertical bars indicate the position of the A+T-rich islands across the *Schizosaccharomyces pombe* chromosomes numbered using a four-digit code, starting with 1001, 2001 and 3001 for chromosomes I, II and III, respectively. The 16 DNA replication origins (ORIs) used to define the properties of A+T-rich islands (Fig. 1A) are indicated by red triangles. The 14 ORIs and autonomous replicating sequence (ARS) elements described by other authors (green triangles) and the 20 islands tested for replication in Fig. 3 (yellow triangles) are also indicated. Only one replication origin at the rRNA cluster is shown (*ars3001*), which coincides with A+T-rich island 3001. Centromeres are represented by a black box, and the A+T-rich islands included in them are labelled red. The *mat* locus in chromosome II is represented by a red box. The asterisk indicates the position of an A+T-rich island in a region missing in the Sanger Centre sequence.

probably due to the relatively large size of the intergenic regions analysed, which made it more likely that they would contain an A+T-rich island. Divergent intergenic regions are longer than the average intergenic distance in *S. pombe* (Wood *et al.*, 2002), and this fact, perhaps in combination with the proximity of two promoters, could contribute to the over-representation of A+T-rich islands in these regions.

A+T-rich islands reliably predict the localization of ORIs

To evaluate the reliability of A+T-rich islands in predicting the localization of ORIs, we selected 20 of them at random (yellow triangles in Fig. 2) and monitored their replication pattern by neutral two-dimensional gel electrophoresis (Brewer *et al.*, 1988; Huberman, 1993). Figure 3A shows that 18 (90%) of the 20 islands tested colocalized with active ORIs, as shown by the presence of intermediates

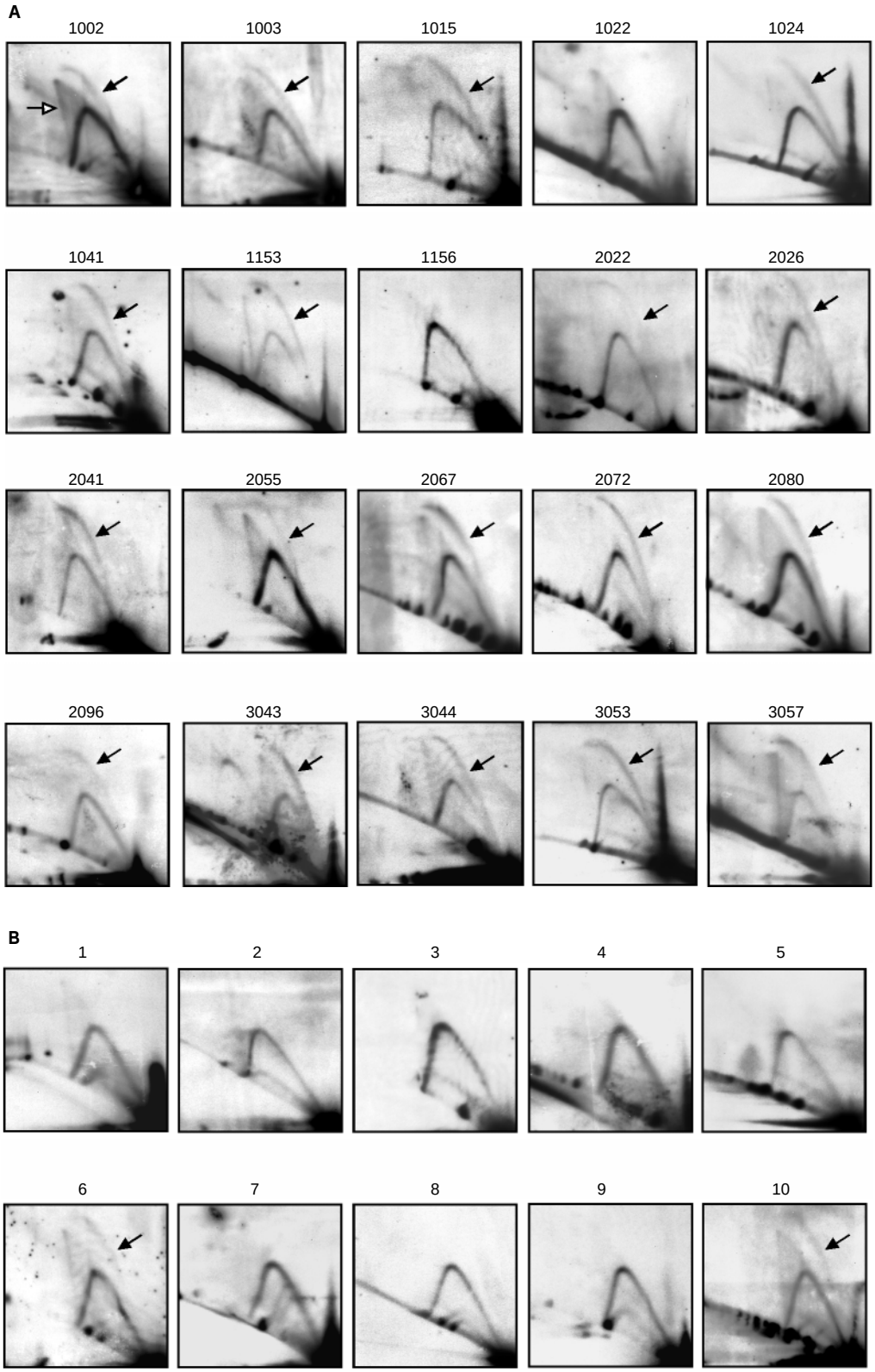

Fig. 3 | A+T-rich islands reliably predict the localization of genomic replication origins. (**A**) Twenty A+T-rich islands (yellow triangles in Fig. 2) were tested for replication by two-dimensional gel electrophoresis. Their identification number in the chromosomes is indicated. Intermediates containing replication bubbles are shown by arrows. A white arrow indicates recombination intermediates. (**B**) Replication analysis of ten genomic regions (1–10) with an A+T content between the intervals defined by the replication origin (ORI)- and non-ORI-containing regions in Fig. 1A.

containing initiation bubbles. This predicts the existence of approximately 345 ORIs associated with A+T-rich islands in the entire genome (in addition to those at the rRNA gene clusters). A spike of recombination intermediates is present in regions containing a replication origin (white arrow in Fig. 3A) as we have previously described (Segurado *et al*., 2002), although the hybridization signal is very weak in a few cases. This could probably be improved by using synchronous cultures (Segurado *et al*., 2002). It is also possible that a small proportion of ORIs that could be active only under certain circumstances or at a certain genomic localization might not be associated with recombination intermediates.

A conspicuous feature of the distribution of A+T-rich islands was their high frequency in the three centromeres, the subtelomeric regions of chromosomes I and II, and the mating-type locus (Fig. 2). Four active ORIs were previously described in the K and L repeats of centromere II (Smith *et al.*, 1995), which turned out to be coincident with islands 2053, 2054, 2057 and 2058. We tested an additional island in centromere II (2055) and two more in the subtelomeric region of chromosome I (1002 and 1003), and in all three cases, they also colocalized with active ORIs. A lower ratio in the bubble to Y arcs in these cases could be due to interference (Dubey et *al.*, 1994), given the high density of predicted ORIs in these regions (Fig. 3A). Island 2068 is immediately adjacent to the *mat1* gene and coincides with the pARS756 sequence (Maundrell *et al.*, 1988) and with an active genomic ORI (Dalgaard & Klar, 2001). The 20-kb region encompassing the *mat2* and *mat3* loci includes A+T-rich islands 2069, 2070 and 2071 plus another one in a 12-kb region missing from the sequence available at the Sanger Centre but included in the National Centre for Biotechnology Information database (labelled with an asterisk in Fig. 2, and not included in table 3 of supplementary information online). Altogether, the average density of A+T-rich islands and ORIs in centromeric, subtelomeric and mating-type regions is about fourfold higher than the genome average.

Because we had defined the A+T-rich islands on the basis of a statistical average, we wondered how strict the boundary was that we had established—that is, how much lower the A+T content of an intergenic region could be relative to the lower limit we have used to define the islands and still act as an ORI. We addressed this by analysing the replication pattern of 18 promoter-containing regions of comparable size to those shown in Fig. 3A but with an A+T content in the 500 bp to 1 kb windows halfway between the intervals defined by ORI and non-ORI regions in Fig. 1A. Analysis of 1 Mb of each chromosome showed that there is an average of 19 intergenic regions per megabase with this base composition, which predicts a total of approximately 240 in the entire genome. We tested 18 of these regions for replication and found that three of them (16.6%) were associated with active ORIs. Bubble arcs were not detected in the other 15, even in overexposed autoradiographs (Fig. 3B and data not shown). The implication is that there could be approximately 40 ORIs in addition to those located at A+T-rich islands, which represents approximately 10% of all ORIs. This also indicates that a small decrease in the A+T content relative to the criterion used to define A+T-rich islands reduces dramatically (from 90% to 16.6%) the reliability in predicting ORIs. Regions with an even lower A+T content probably contain very few ORIs, if any. This is suggested by the fact that none of the 16 regions with an A+T content close to the intergenic genomic average, was associated with ORIs (Fig. 1A, white circles) and by the localization of all the previously identified ORIs with A+T-rich islands. Altogether, our results indicate that A+T-rich islands are very reliable predictors of ORIs in *S. pombe* and that the proportion of ORIs not associated with them is likely to be small.

DISCUSSION

This work establishes A+T-rich islands as discrete genomic regions that can account for the distinctive properties of the *S. pombe* ORIs. Recent studies have suggested that several ORC binding sites are collectively required for efficient ORI firing (Kong & DePamphilis, 2002; Takahashi *et al.*, 2003). Cooperation between ORC complexes to attain the critical concentration to trigger replication would require a minimal length of A+T-rich DNA, which is consistent with the size of the A+T-rich islands and with the 0.5–1 kb length of *S. pombe* ORIs previously described. Cooperativity and a certain degree of redundancy are also suggested by the fact that progressive shortening of several ORI-containing regions results in a gradual decline in replication efficiency, rather than in an all-or-none effect (Dubey *et al.*, 1994; Clyne & Kelly, 1995; Okuno *et al.*, 1999). Redundancy can also explain that removal of a 330-bp-long region encompassing the replication initiation point in *ars1* in the chromosome diminishes, but does not prevent, ORI activity (Gómez & Antequera, 1999). One of the best-characterized ORIs in *S. pombe* is *ars2004* (Okuno *et al.*, 1997, 1999; Takahashi *et al*., 2003) in which three specific regions have been shown to be collectively essential for ORI activity in the chromosome (Takahashi *et al*., 2003). Region I includes a tract of poly-adenine 19 bp long, and region III contains 11 repeats of the AAAAT sequence. These elements, however, are not present in many other *S. pombe* ORIs, suggesting that, although important for *ars2004*, they are not a general requirement. The lack of conserved sequence elements between A+T-rich islands suggests that the SpOrc4 protein can bind, via its AT-hooks, A+T-rich sequences of different composition at different ORIs.

Our data predict the existence of approximately 385 ORIs in the *S. pombe* genome, of which 345 would be located at A+T-rich islands and another 40 at regions with a slightly lower A+T content (Fig. 3). Although it is possible that there could be a few additional ORIs at other genomic regions, our estimate is very close to the recent prediction of approximately 400 ORIs in *S. cerevisiae* (Raghuraman *et al.*, 2001; Wyrick *et al.*, 2001). However, no correlation between ORIs and a distinctive base composition has been found in *S. cerevisiae* (Raghuraman *et al.*, 2001). The distribution of A+T-rich islands in *S. pombe* is relatively homogeneous except at centromeres, subtelomeric regions and the mating-type locus, where the average frequency of A+T-rich islands is one every 8 kb. A high density of ORIs has also been found in many subtelomeric regions in *S. cerevisiae* (Wyrick *et al.*, 2001). This suggests a possible role for ORC (or for other components of the replication initiation machinery) in establishing or maintaining the chromatin organization and transcriptional silence in these regions in *S. pombe*, as has been described in *S. cerevisiae* (Fox *et al.*, 1995; Palacios DeBeer *et al.,* 2003) and *Drosophila* (Pak *et al.*, 1997). The chromoprotein Swi6 could be involved in this link given that it binds to silent chromatin in centromeres, telomeres and the mating-type locus and interacts directly with DNA polymerase-α (Ahmed *et al*., 2001; Nakayama *et al*., 2001). *S. pombe* ORIs are similar to those of mammalian cells in their lack of consensus sequences and in their preference to localize in intergenic regions close to promoters. A striking difference, however, is that many replication origins in mammals are associated with G+C-rich islands (Delgado *et al.*, 1998),

which are regions with a G+C content that is higher than the genome average. It is possible that replication initiation is associated with some kind of genetic instability that, along the evolution of some organisms, could have shifted the base composition of ORI regions in two alternative directions to generate either G+C-rich or A+T-rich islands.

METHODS

Base composition analysis. This study was performed taking as a reference the Sep-05-2002 version of the *S. pombe* genome sequence available at the Sanger Centre. The telomeric repeats and the 1.1 Mb rRNA gene clusters were not included in the analysis. We scanned the entire sequence with the base composition tool of the Artemis software package (http://www.sanger.ac.uk/Software/Artemis) using windows of 500 bp to 1 kb and a step of 1 bp. The regions that had an A+T content above the limits indicated in the text were considered A+T-rich islands. Their position in the genome is indicated in table 3 of supplementary information online. For the analysis of specific regions (Fig. 1B,C), the corresponding sequences were downloaded and scanned with the FREAK program of the EMBOSS software package (http://www.emboss.org) using a 500 bp window and a step of 100 bp.

Culture conditions and two-dimensional gel electrophoresis analysis. Cultures of *S. pombe* h$^-$972 grown in rich medium were used for all the experiments. DNA from 500 ml of an exponentially growing culture (A_{595} = 0.8) was isolated and used for each gel as previously described (Segurado *et al.*, 2002). Replication intermediates were separated by two-dimensional neutral gel electrophoresis under conditions previously described (Brewer *et al.*, 1988; Huberman, 1993). Specific genomic regions were selected for two-dimensional gel electrophoresis according to the availability of restriction sites that would give rise to a restriction fragment between 3 kb and 6 kb long, which is the optimal size range required for this analysis. The position of restriction fragments and probes used for the 38 regions tested (Fig. 3) are available on request.

Supplementary information is available at EMBO reports online (http://www.emboreports.org).

ACKNOWLEDGEMENTS

We are grateful to A. Aguilera, J. Casadesús, M. Gómez, P. Hernández and M. Tamame for many suggestions and for helpful comments on the manuscript. We also thank V. Wood for the data on the distribution of intergenic regions. M.S. and A.deL. were supported by postgraduate fellowships from the Spanish Ministerio de Ciencia y Tecnología (MCYT) and this work was funded by grant BMC2002-03591 (from the MCYT).

REFERENCES

Ahmed, S., Saini, S., Arora, S. & Singh, J. (2001) Chromodomain protein Swi6-mediated role of DNA polymerase α in establishment of silencing in fission yeast. *J. Biol. Chem.*, **276**, 47814–47821.

Brewer, B.J., Sena, E. P. & Fangman, W.L. (1988) in *Cancer Cells* Vol. 6 (eds Kelly, T. & Stillman, B.) 229–234. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

Broach, J.R., *et al*. (1983) Localization and sequence analysis of yeast origins of DNA replication. *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 1165–1173.

Chuang, R.Y. & Kelly, T.J. (1999) The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proc. Natl Acad. Sci. USA*, **96**, 2656–2661.

Clyne, R.K. & Kelly, T.J. (1995) Genetic analysis of an ARS element from the fission yeast *Schizosaccharomyces pombe*. *EMBO J.*, **14**, 6348–6357.

Dalgaard, J.Z. & Klar, A.J. (2001) A DNA replication-arrest site RTS1 regulates imprinting by determining the direction of replication at mat1 in *S. pombe*. *Genes Dev.*, **15**, 2060–2068.

Delgado, S., Gómez, M., Bird, A. & Antequera, F. (1998) Initiation of DNA replication at CpG islands in mammalian chromosomes. *EMBO J.* **17**, 2426–2435.

Dubey, D.D., Zhu, J., Carlson, D.L. Sharma, K. & Huberman, J.A. (1994) Three ARS elements contribute to the ura4 replication origin region in the fission yeast *Schizosaccharomyces pombe*. *EMBO J.*, **13**, 3638–3647.

Dubey, D.D., Kim, S.M., Todorov, I.T. & Huberman, J.A. (1996) Large, complex modular structure of a fission yeast DNA replication origin. *Curr. Biol.*, **6**, 467–473.

Fox, C.A., Loo, S., Dillin, A. & Rine, J. (1995) The origin recognition complex has essential functions in transcriptional silencing and chromosomal replication. *Genes Dev.*, **9**, 911–924.

Gómez, M. & Antequera, F. (1999) Organization of DNA replication origins in the fission yeast genome. *EMBO J.*, **18**, 5683–5690.

Huberman, J.A. (1993) in *The Cell Cycle. A Practical Approach* (eds Fantes, P. & Brooks, R.) 213–234. Oxford University Press, Oxford, UK.

Kong, D. & DePamphilis, M.L. (2002) Site-specific ORC binding, pre-replication complex assembly and DNA synthesis at *Schizosaccharomyces pombe* replication origins. *EMBO J.*, **21**, 5567–5576.

Lee, J.K., Moon, K.Y., Jiang, Y. & Hurwitz, J. (2001) The *Schizosaccharomyces pombe* origin recognition complex interacts with multiple AT-rich regions of the replication origin DNA by means of the AT-hook domains of the spOrc4 protein. *Proc. Natl Acad. Sci. USA*, **98**, 13589–13594.

Maundrell, K., Hutchinson, A. & Shall, S. (1988) Sequence analysis of ARS elements in fission yeast. *EMBO J.*, **7**, 2203–2209.

Nakayama, J., Allshire, R.C., Klar, A.J. & Grewal, S.I. (2001) A role for DNA polymerase α in epigenetic control of transcriptional silencing in fission yeast. *EMBO J.*, **20**, 2857–2866.

Newlon, C.S. & Theis, J.F. (1993) The structure and function of yeast ARS elements. *Curr. Opin. Gen. Dev.* **3**, 752–758.

Okuno, Y., Okazaki, T. & Masukata, H. (1997) Identification of a predominant replication origin in fission yeast. *Nucleic Acids Res.*, **25**, 530–536.

Okuno, Y., Satoh, H., Sekiguchi, M. & Masukata, H. (1999) Clustered adenine/thymine stretches are essential for function of a fission yeast replication origin. *Mol. Cell. Biol.*, **19**, 6699–6709.

Pak, D.T., *et al*. (1997) Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes. *Cell*, **91**, 311–323.

Palacios DeBeer, M.A., Müller, U. & Fox, C.A. (2003) Differential DNA affinity specifies roles for the origin recognition complex in budding yeast heterochromatin. *Genes Dev.* **17**, 1817–1822.

Raghuraman, M.K. *et al*. (2001) Replication dynamics of the yeast genome. *Science*, **294**, 115–121.

Reeves, R. & Beckerbauer, L. (2001) HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim. Biophys. Acta*, **1519**, 13–29.

Sanchez, A., Kim, S.M. & Huberman, J. (1998) Ribosomal DNA replication in the fission yeast *Schizosaccharomyces pombe*. *Exp. Cell Res.* **238**, 220–230.

Segurado, M., Gómez, M. & Antequera, F. (2002) Increased recombination intermediates and homologous integration hot spots at DNA replication origins. *Mol. Cell*, **10**, 907–916.

Smith, J.G. *et al*. (1995) Replication of centromere II of *Schizosaccharomyces pombe*. *Mol. Cell. Biol.*, **15**, 5165–5172.

Takahashi, T., Ohara, E., Nishitani, H. & Masukata, H. (2003) Multiple ORC binding sites are required for efficient MCM loading and origin firing in fission yeast. *EMBO J.*, **22**, 964–974.

Theis, J.F. & Newlon, C.S. (1997) The ARS309 chromosomal replicator *of Saccharomyces cerevisiae* depends on an exceptional ARS consensus sequence. *Proc. Natl Acad. Sci. USA*, **94**, 10786–10791.

Theis, J.F. & Newlon, C.S. (2001) Two compound replication origins in *Saccharomyces cerevisiae* contain redundant origin recognition complex binding sites. *Mol. Cell. Biol.*, **21**, 2790–2801.

Wood, V. *et al.* (2002) The genome sequence of *Schizosaccharomyces pombe*. *Nature*, **415**, 871–880.

Wyrick, J. *et al*. (2001) Genome-wide distribution of ORC and MCM proteins in *S. cerevisiae*: high-resolution mapping of replication origins. *Science*, **294**, 2357–2360.